Supplementary Material to "Characterizing the Risk of Atrial Fibrillation in Cardiac Patients with Exceptional Electrocardiogram Phenotypes"

Lieke van den Biggelaar
l.a.j.v.d.biggelaar@tue.nl
Eindhoven University of Technology
Eindhoven, the Netherlands

Rianne M. Schouten
r.m.schouten@tue.nl
Eindhoven University of Technology
Eindhoven, the Netherlands

Ashley de Bie
ashley.d.bie@catharinaziekenhuis.nl
Catharina Hospital
Eindhoven, the Netherlands

R. Arthur Bouwman
arthur.bouwman@catharinaziekenhuis.nl
Catharina Hospital
Eindhoven, the Netherlands

Wouter Duivesteijn
w.duivesteijn@tue.nl
Eindhoven University of Technology
Eindhoven, the Netherlands

Abstract

As a supplement to our paper [1], we provide two sections. On the one hand, we provide the full tables of subgroups discovered from all eleven experiments listed in [1, Table 3]. On the other hand, we provide additional experiments on publicly available data, to mirror the experiments in [1, Section 6] which are run on proprietary hospital data, in a way that makes the results in this document reproducible.

1 Additional Results on Hospital Data

Having completed the eleven experiments outlined in [1, Table 3], we reported only the medically most relevant findings (as discussed with medical doctors at the hospital) in [1, Section 6.3]. Here, we list for each experiment the full top-15 subgroups reported by beam search for each of the eleven experiments. Tables 1–11 list these results for the experiments 1–11 from [1, Table 3], respectively.

Table 1: Top-15 subgroups discovered when seeking exceptional phenotype θ_{SDSD}.

Phenotype	#	Description
θ_{SDSD}	1	anion gap venous $\in (0.0; 3.0) \wedge$ age $\in (70; 88) \wedge$ fluid loss $\in (0.0; 0.0)$
θ_{SDSD}	2	age $\in (71; 88) \wedge$ anion gap venous $\in (0.0; 3.0) \wedge$ fluid loss $\in (0.0; 0.0)$
θ_{SDSD}	3	age $\in (71; 88) \wedge$ Euroscore-I $\in (5.7; 40.7) \wedge$ potassium $\in (3.9; 5.1)$
θ_{SDSD}	4	glucose (arterial) $\in (4.0; 6.0) \wedge$ blood loss $\in (641.6; 3868.0) \wedge$ Midazolam-hameln 5 mg/ml IV administration $\in (0.0; 0.0)$
θ_{SDSD}	5	Euroscore-I $\in (6.4; 85.7) \wedge$ age $\in (71; 88) \wedge$ potassium $\in (4.1; 5.1)$
θ_{SDSD}	6	glucose (arterial) $\in (4.0; 6.0) \wedge$ blood loss $\in (641.6; 3868.0) \wedge$ anion gap $\in (0.0; 6.0)$
θ_{SDSD}	7	anion gap venous $\in (0.0; 3.0) \wedge$ age $\in (70; 88) \wedge$ lactate $\in (1.3; 2.8)$
θ_{SDSD}	8	age $\in (71; 88) \wedge$ anion gap venous $\in (0.0; 3.0) \wedge$ lactate $\in (1.3; 2.8)$
θ_{SDSD}	9	anion gap venous $\in (0.0; 3.0) \wedge$ age $\in (70; 88) \wedge$ fluid loss $\in (0.0; 9.6)$
θ_{SDSD}	10	age $\in (71; 88) \wedge$ anion gap venous $\in (0.0; 3.0) \wedge$ fluid loss $\in (0.0; 9.6)$
θ_{SDSD}	11	anion gap venous $\in (0.0; 3.0) \wedge$ age $\in (73; 88) \wedge$ lactate $\in (1.3; 2.8)$
θ_{SDSD}	12	age $\in (75; 88) \wedge$ protamin administration $\in (14.3; 21.1) \wedge$ PCO2 $\in (37.0; 45.2)$
θ_{SDSD}	13	glucose arterial $\in (4.0; 6.0) \wedge$ blood loss $\in (641.6; 3868.0) \wedge$ SO2 $\in (77.8; 100.0)$
θ_{SDSD}	14	Euroscore-I $\in (2.4; 85.7) \wedge$ taking pantoprazole = True $\wedge$ priority $\in$ {unknown, plannable, < 1 week}
θ_{SDSD}	15	Euroscore-I $\in (2.4; 85.7) \wedge$ taking pantoprazole = True $\wedge$ priority $\in$ {unknown, plannable, < 1 week, < 72 hours}

Table 2: Top-15 subgroups discovered when seeking exceptional phenotype θ_{RMSSD}.

Phenotype	#	Description
θ_{RMSSD}	1	anion gap venous $\in (0.0; 3.0) \wedge$ age $\in (70; 88) \wedge$ fluid loss $\in (0.0; 0.0)$
θ_{RMSSD}	2	age $\in (71; 88) \wedge$ anion gap venous $\in (0.0; 3.0) \wedge$ fluid loss $\in (0.0; 0.0)$
θ_{RMSSD}	3	anion gap venous $\in (0.0; 3.0) \wedge$ age $\in (73; 88) \wedge$ potassium $\in (1.3; 2.8)$
θ_{RMSSD}	4	anion gap venous $\in (0.0; 3.0) \wedge$ age $\in (70; 88) \wedge$ lactate $\in (1.3; 2.8)$
θ_{RMSSD}	5	age $\in (71; 88) \wedge$ anion gap venous $\in (0.0; 3.0) \wedge$ lactate $\in (1.3; 2.8)$
θ_{RMSSD}	6	anion gap venous $\in (0.0; 3.0) \wedge$ age $\in (70; 88) \wedge$ fluid loss $\in (0.0; 9.6)$
θ_{RMSSD}	7	age $\in (71; 88) \wedge$ anion gap venous $\in (0.0; 3.0) \wedge$ fluid loss $\in (0.0; 9.6)$
θ_{RMSSD}	8	glucose (arterial) $\in (4.0; 6.0) \wedge$ blood loss $\in (641.6; 3868.0) \wedge$ thrombocytes $\in (99.0; 170.0)$
θ_{RMSSD}	9	glucose (arterial) $\in (4.0; 6.0) \wedge$ blood loss $\in (641.6; 3868.0) \wedge$ anion gap $\in (0.0; 6.0)$
θ_{RMSSD}	10	anion gap venous $\in (0.0; 3.0) \wedge$ alfentanil administration $\in (3.0; 21.3) \wedge$ taking atorvastatin = False
θ_{RMSSD}	11	anion gap venous $\in (0.0; 3.0) \wedge$ alfentanil administration $\in (3.0; 21.3)$
θ_{RMSSD}	12	milrinone administration $\in (3.5; 635.4) \wedge$ phenylephrine administration $\in (0.0; 9.4) \wedge$ fibrinogen $\in (2.1; 4.4)$
θ_{RMSSD}	13	glucose (arterial) $\in (4.0; 6.0) \wedge$ blood loss $\in (641.6; 3868.0) \wedge$ Midazolam-hameln 5 mg/ml IV administration $\in (0.0; 0.0)$
θ_{RMSSD}	14	anion gap venous $\in (0.0; 3.0) \wedge$ age $\in (73; 88) \wedge$ fluid loss $\in (0.0; 10.0)$
θ_{RMSSD}	15	age $\in (75; 88) \wedge$ Euroscore-I $\in (2.4; 85.7) \wedge$ priority $\in$ {unknown, plannable, $<$ 1 week, $<$ 72 hours}

Table 3: Top-15 subgroups discovered when seeking exceptional phenotype θ_{SDRR}.

Phenotype	#	Description
θ_{SDRR}	1	age $\in (71; 88) \wedge$ taking acetylsalicylic acid = True $\wedge$ morphine administration $\in (0.0; 0.9)$
θ_{SDRR}	2	anion gap venous $\in (0.0; 3.0) \wedge$ age $\in (70; 88) \wedge$ fluid loss $\in (0.0; 0.0)$
θ_{SDRR}	3	age $\in (71; 88) \wedge$ anion gap venous $\in (0.0; 3.0) \wedge$ fluid loss $\in (0.0; 0.0)$
θ_{SDRR}	4	anion gap venous $\in (0.0; 3.0) \wedge$ APTTCkprest = unknown $\wedge$ hemoglobin $\in (3.7; 6.5)$
θ_{SDRR}	5	age $\in (71; 88) \wedge$ bicarbonate $\in (19.0; 23.6) \wedge$ NovoRapid administration $\in (0.0; 0.0)$
θ_{SDRR}	6	glucose $\in (4.1; 6.0) \wedge$ blood loss $\in (570.4; 3868.0) \wedge$ Midazolam-hameln 5 mg/ml IV administration $\in (0.0; 0.0)$
θ_{SDRR}	7	glucose (arterial) $\in (4.0; 6.0) \wedge$ blood loss $\in (641.6; 3868.0) \wedge$ SO2 $\in (77.8; 100.0)$
θ_{SDRR}	8	anion gap venous $\in (0.0; 3.0) \wedge$ age $\in (70; 88) \wedge$ fluid loss $\in (0.0; 9.6)$
θ_{SDRR}	9	age $\in (71; 88) \wedge$ anion gap venous $\in (0.0; 3.0) \wedge$ fluid loss $\in (0.0; 9.6)$
θ_{SDRR}	10	glucose (arterial) $\in (4.0; 6.0) \wedge$ alfentanil administration $\in (29.4; 54.6) \wedge$ calcium gluconate administration $\in (20.0; 70.0)$
θ_{SDRR}	11	anion gap venous $\in (0.0; 3.0) \wedge$ APTTCkprest = unknown $\wedge$ carboxyhemoglobin $\in (1.2; 1.8)$
θ_{SDRR}	12	age $\in (71; 88) \wedge$ bicarbonate $\in (19.0; 23.6) \wedge$ Midazolam-hameln 5 mg/ml IV administration $\in (0.0; 0.0)$
θ_{SDRR}	13	age $\in (75; 88) \wedge$ bicarbonate $\in (19.6; 25.6) \wedge$ smoking $\in$ {never done, done in the past}
θ_{SDRR}	14	anion gap venous $\in (0.0; 3.0) \wedge$ age $\in (70; 88) \wedge$ lactate $\in (1.3; 2.8)$
θ_{SDRR}	15	age $\in (71; 88) \wedge$ anion gap venous $\in (0.0; 3.0) \wedge$ lactate $\in (1.3; 2.8)$

Table 4: Top-15 subgroups discovered when seeking exceptional phenotype θ_P.

Phenotype	#	Description
θ_P	1	sodium $\in (112.0; 135.0) \wedge$ taking metoprolol = True $\wedge$ fluid loss $\in (0.0; 24.0)$
θ_P	2	sodium $\in (112.0; 135.0) \wedge$ taking metoprolol = True $\wedge$ fluid loss $\in (0.0; 0.0)$
θ_P	3	sodium $\in (112.0; 135.0) \wedge$ taking metoprolol = True $\wedge$ coronary artery bypass grafting = False
θ_P	4	creatinine $\in (48.0; 73.6) \wedge$ PO2 $\in (38.0; 230.0) \wedge$ dexamethasone administration $\in (0.0; 4.1)$
θ_P	5	Euroscore-I $\in (2.4; 85.7) \wedge$ CK $\in (303.0; 841.0) \wedge$ phenylephrine administration $\in (0.0; 3.5)$
θ_P	6	Euroscore-I $\in (2.4; 85.7) \wedge$ CK $\in (303.0; 841.0) \wedge$ methemoglobin $\in (0.7; 1.9)$
θ_P	7	cardioplegia $\in (1078.4; 4629.0) \wedge$ base excess $\in (-1.3; 3.4) \wedge$ glucose $\in (5.2; 7.7)$
θ_P	8	hemoglobin $\in (3.7; 5.8) \wedge$ eGFR (CKD-EPI) $\in (62.4; 91.0) \wedge$ anion gap $\in (0.0; 4.0)$
θ_P	9	hemoglobin $\in (3.7; 5.8) \wedge$ eGFR (CKD-EPI) $\in (62.4; 91.0) \wedge$ carboxyhemoglobin $\in (0.9; 1.8)$
θ_P	10	taking pantoprazole = True $\wedge$ thrombocytes $\in (120.0; 244.0) \wedge$ morphine administration $\in (0.0; 0.9)$
θ_P	11	hemoglobin $\in (3.7; 5.8) \wedge$ creatinine $\in (62.0; 98.0) \wedge$ age $\in (54; 73)$
θ_P	12	Euroscore-I $\in (2.4; 85.7) \wedge$ CK $\in (303.0; 841.0) \wedge$ methemoglobin $\in (0.8; 1.9)$
θ_P	13	CK $\in (362.2; 1723.0) \wedge$ Euroscore-I $\in (1.6; 16.6) \wedge$ glucose $\in (6.2; 11.4)$
θ_P	14	CK $\in (362.2; 1723.0) \wedge$ Euroscore-I $\in (1.6; 16.6) \wedge$ age $\in (61; 73)$
θ_P	15	cardioplegia $\in (1078.4; 4629.0) \wedge$ base excess $\in (-1.3; 3.4) \wedge$ morphine administration $\in (0.0; 0.8)$

Table 5: Top-15 subgroups discovered when seeking exceptional phenotype θ_P.

Phenotype	#	Description
θ_F	1	hematocrit $\in (0.2; 0.3) \wedge$ glucose $\in (8.2; 12.7) \wedge$ drug usage $\in$ {never used, used in the past}
θ_F	2	hematocrit $\in (0.2; 0.3) \wedge$ glucose $\in (8.2; 12.7) \wedge$ drug usage $\in$ {never used}
θ_F	3	hematocrit $\in (0.2; 0.3) \wedge$ glucose $\in (8.2; 12.7) \wedge$ albumin administration $\in (0.0; 0.0)$
θ_F	4	hematocrit $\in (0.2; 0.3) \wedge$ glucose $\in (8.2; 12.7) \wedge$ modified gelatin administration $\in (0.0; 0.0)$
θ_F	5	neutrophils $\in (6.8; 23.0) \wedge$ CK $\in (91.0; 554.8) \wedge$ morphine administration $\in (0.2; 2.8)$
θ_F	6	neutrophils $\in (6.8; 23.0) \wedge$ morphine administration $\in (0.2; 2.8) \wedge$ PO2 $\in (38.0; 196.4)$
θ_F	7	neutrophils $\in (6.8; 23.0) \wedge$ CK $\in (91.0; 554.8) \wedge$ cell salvage administration $\in (0.0; 1318.3)$
θ_F	8	neutrophils $\in (6.8; 23.0) \wedge$ PO2 $\in (36.0; 194.2) \wedge$ clonidine administration $\in (0.0; 0.0)$
θ_F	9	eosinophils $\in (-1.0; -1.0) \wedge$ potassium $\in (4.2; 5.8) \wedge$ PO2 $\in (38.0; 192.0)$
θ_F	10	APTT (heparanase) $\in (27.4; 46.5) \wedge$ hematocrit $\in (0.24; 0.33) \wedge$ BMI $\in$ {underweight, healthy, overweight, obese}
θ_F	11	leucocytes $\in (8.1; 57.6) \wedge$ monocytes $\in (0.2; 0.8) \wedge$ taking losartan with diuretics = False
θ_F	12	APTT (heparanase) $\in (27.4; 46.5) \wedge$ hematocrit $\in (0.2; 0.3) \wedge$ pH $\in (7.3; 7.4)$
θ_F	13	hematocrit $\in (0.2; 0.3) \wedge$ rocuronium administration $\in (0.0; 0.0) \wedge$ priority $\in$ {unknown, plannable, $<$ 1 week, $<$ 72 hours}
θ_F	14	neutrophils $\in (6.8; 23.0) \wedge$ CK $\in (91.0; 554.8) \wedge$ thrombocytes $\in (67.0; 232.0)$
θ_F	15	neutrophils $\in (6.8; 23.0) \wedge$ morphine administration $\in (0.2; 2.8) \wedge$ potassium $\in (3.8; 5.8)$

Table 6: Top-15 subgroups discovered when seeking exceptional phenotype $\theta^P_{\mathbf{SDSD}}$.

Phenotype	#	Description
θ^P_{SDSD}	1	prothrombin time $\in (16.5; 131.0) \wedge$ standard bicarbonate $\in (23.7; 27.3) \wedge$ thrombin administration $\in (0.0; 0.0)$
θ^P_{SDSD}	2	fibrinogen $\in (1.7; 7.7) \wedge$ base excess $\in (-0.6; 7.3) \wedge$ thrombin administration $\in (0.0; 0.0)$
θ^P_{SDSD}	3	prothrombin time $\in (17.7; 131.0) \wedge$ base excess $\in (-0.6; 3.2) \wedge$ calcium ions $\in (1.1; 1.2)$
θ^P_{SDSD}	4	milrinone administration $\in (3.5; 635.4) \wedge$ dexamethasone administration $\in (0.0; 2.6) \wedge$ chloride $\in (97.0; 111.0)$
θ^P_{SDSD}	5	anion gap $\in (-1.0; 2.0) \wedge$ glucose (arterial) $\in (4.7; 6.7) \wedge$ minimally invasive aortic valve replacement = False
θ^P_{SDSD}	6	anion gap $\in (-1.0; 2.0) \wedge$ glucose $\in (4.1; 6.7) \wedge$ sex = male
θ^P_{SDSD}	7	prothrombin time $\in (17.7; 131.0) \wedge$ noradrenalin administration $\in (0.0; 24.6) \wedge$ fluid loss $\in (0.0; 0.0)$
θ^P_{SDSD}	8	prothrombin time $\in (17.7; 131.0) \wedge$ blood loss $\in (0.0; 958.8) \wedge$ age $\in (51; 74)$
θ^P_{SDSD}	9	prothrombin time $\in (16.5; 131.0) \wedge$ standard bicarbonate $\in (23.7; 27.3) \wedge$ chloride $\in (105.0; 110.0)$
θ^P_{SDSD}	10	prothrombin time $\in (17.7; 131.0) \wedge$ standard bicarbonate $\in (23.8; 27.3) \wedge$ calcium ions $\in (1.1; 1.2)$
θ^P_{SDSD}	11	prothrombin time $\in (17.7; 131.0) \wedge$ standard bicarbonate $\in (23.8; 27.3) \wedge$ sodium $\in (133.0; 141.0)$
θ^P_{SDSD}	12	prothrombin time $\in (17.7; 131.0) \wedge$ standard bicarbonate $\in (23.8; 27.3) \wedge$ alfentanil administration $\in (0.0; 42.3)$
θ^P_{SDSD}	13	prothrombin time $\in (17.7; 131.0) \wedge$ blood loss $\in (0.0; 958.8) \wedge$ fluid loss $\in (0.0; 0.0)$
θ^P_{SDSD}	14	prothrombin time $\in (16.5; 131.0) \wedge$ standard bicarbonate $\in (23.7; 27.3) \wedge$ sodium $\in (132.2; 141.0)$
θ^P_{SDSD}	15	milrinone administration $\in (3.5; 635.4) \wedge$ cefazolin administration $\in (0.0; 92.5) \wedge$ prothrombin time $\in (14.8; 131.0)$

Table 7: Top-15 subgroups discovered when seeking exceptional phenotype $\theta^P_{\mathbf{RMSSD}}$.

Phenotype	#	Description
θ^P_{RMSSD}	1	prothrombin time $\in (16.5; 131.0) \wedge$ standard bicarbonate $\in (23.7; 27.3) \wedge$ thrombin administration $\in (0.0; 0.0)$
θ^P_{RMSSD}	2	fibrinogen $\in (1.7; 7.7) \wedge$ base excess $\in (-0.6; 7.3) \wedge$ thrombin administration $\in (0.0; 0.0)$
θ^P_{RMSSD}	3	prothrombin time $\in (17.7; 131.0) \wedge$ standard bicarbonate $\in (23.8; 27.3) \wedge$ alfentanil administration $\in (0.0; 42.3)$
θ^P_{RMSSD}	4	prothrombin time $\in (17.7; 131.0) \wedge$ base excess $\in (-0.7; 3.2) \wedge$ calcium ions $\in (1.1; 1.2)$
θ^P_{RMSSD}	5	prothrombin time $\in (17.7; 131.0) \wedge$ cefazolin administration $\in (0.0; 90.3) \wedge$ FIO2 $\in (0.2; 0.2)$
θ^P_{RMSSD}	6	prothrombin time $\in (17.7; 131.0) \wedge$ standard bicarbonate $\in (23.8; 27.3) \wedge$ calcium ions $\in (1.1; 1.2)$
θ^P_{RMSSD}	7	prothrombin time $\in (16.5; 131.0) \wedge$ standard bicarbonate $\in (23.7; 27.3) \wedge$ chloride $\in (105.0; 110.0)$
θ^P_{RMSSD}	8	prothrombin time $\in (17.7; 131.0) \wedge$ standard bicarbonate $\in (23.8; 27.3) \wedge$ sodium $\in (133.0; 141.0)$
θ^P_{RMSSD}	9	prothrombin time $\in (17.7; 131.0) \wedge$ blood loss $\in (0.0; 958.8) \wedge$ age $\in (51; 74)$
θ^P_{RMSSD}	10	milrinone administration $\in (3.5; 635.4) \wedge$ cefazolin administration $\in (0.0; 92.5) \wedge$ prothrombin time $\in (14.8; 131.0)$
θ^P_{RMSSD}	11	prothrombin time $\in (17.7; 131.0) \wedge$ noradrenalin administration $\in (0.0; 24.6) \wedge$ fluid loss $\in (0.0; 0.0)$
θ^P_{RMSSD}	12	prothrombin time $\in (17.7; 131.0) \wedge$ blood loss $\in (0.0; 958.8) \wedge$ fluid loss $\in (0.0; 0.0)$
θ^P_{RMSSD}	13	prothrombin time $\in (17.7; 131.0) \wedge$ blood loss $\in (0.0; 958.8) \wedge$ chloride $\in (105.0; 111.0)$
θ^P_{RMSSD}	14	fibrinogen $\in (1.7; 7.7) \wedge$ standard bicarbonate $\in (23.8; 30.2) \wedge$ drains $\in (0.0; 480.0)$
θ^P_{RMSSD}	15	prothrombin time $\in (17.7; 131.0) \wedge$ base excess $\in (-0.7; 3.2) \wedge$ coronary artery bypass grafting = False

Table 8: Top-15 subgroups discovered when seeking exceptional phenotype θ^P_{SDSQ}.

Phenotype	#	Description
θ^P_{SDSQ}	1	prothrombin time $\in (17.7; 131.0) \wedge$ standard bicarbonate $\in (23.8; 27.3) \wedge$ age $\in (51; 74)$
θ^P_{SDSQ}	2	prothrombin time $\in (17.7; 131.0) \wedge$ noradrenalin administration $\in (0.0; 24.6) \wedge$ fluid loss $\in (0.0; 0.0)$
θ^P_{SDSQ}	3	fibrinogen $\in (1.7; 7.7) \wedge$ base excess $\in (-0.6; 7.3) \wedge$ thrombin administration $\in (0.0; 0.0)$
θ^P_{SDSQ}	4	base excess $\in (-0.6; 7.3) \wedge$ fibrinogen $\in (2.1; 7.7) \wedge$ thrombin administration $\in (0.0; 0.0)$
θ^P_{SDSQ}	5	anion gap $\in (-1.0; 2.0) \wedge$ glucose $\in (4.1; 6.7) \wedge$ sex = male
θ^P_{SDSQ}	6	prothrombin time $\in (17.7; 131.0) \wedge$ blood loss $\in (0.0; 958.8) \wedge$ age $\in (51; 74)$
θ^P_{SDSQ}	7	prothrombin time $\in (17.7; 131.0) \wedge$ cefazolin administration $\in (0.0; 90.3) \wedge$ taking pantoprazole = False
θ^P_{SDSQ}	8	standard bicarbonate $\in (23.7; 30.2) \wedge$ alfentanil administration $\in (30.5; 55.4) \wedge$ calcium ions $\in (1.1; 1.2)$
θ^P_{SDSQ}	9	prothrombin time $\in (17.7; 131.0) \wedge$ blood loss $\in (0.0; 958.8) \wedge$ fluid loss $\in (0.0; 0.0)$
θ^P_{SDSQ}	10	base excess $\in (-0.6; 7.3) \wedge$ prothrombin time $\in (16.4; 131.0) \wedge$ alfentanil administration $\in (0.0; 46.3)$
θ^P_{SDSQ}	11	base excess $\in (-0.6; 7.3) \wedge$ prothrombin time $\in (16.4; 131.0) \wedge$ fibrinogen administration $\in (0.0; 40.0)$
θ^P_{SDSQ}	12	base excess $\in (-0.6; 7.3) \wedge$ prothrombin time $\in (16.4; 131.0) \wedge$ fibrinogen administration $\in (0.0; 0.0)$
θ^P_{SDSQ}	13	prothrombin time $\in (17.7; 131.0) \wedge$ base excess $\in (-0.7; 3.2) \wedge$ calcium ions $\in (1.1; 1.2)$
θ^P_{SDSQ}	14	glucose (arterial) $\in (4.0; 6.7) \wedge$ anion gap $\in (-1.0; 2.0) \wedge$ minimally invasive aortic valve replacement = False
θ^P_{SDSQ}	15	base excess $\in (-0.6; 7.3) \wedge$ prothrombin time $\in (16.4; 131.0) \wedge$ age $\in (51; 73)$

Table 9: Top-15 subgroups discovered when seeking exceptional phenotype θ^F_{SDSD}.

Phenotype	#	Description
θ^F_{SDSD}	1	prothrombin time $\in (16.5; 131.0) \wedge$ standard bicarbonate $\in (23.7; 27.3) \wedge$ thrombin administration $\in (0.0; 0.0)$
θ^F_{SDSD}	2	prothrombin time $\in (17.7; 131.0) \wedge$ cefazolin administration $\in (0.0; 90.3) \wedge$ FIO2 $\in (0.2; 0.2)$
θ^F_{SDSD}	3	prothrombin time $\in (17.7; 131.0) \wedge$ noradrenalin administration $\in (0.0; 24.6) \wedge$ fluid loss $\in (0.0; 0.0)$
θ^F_{SDSD}	4	prothrombin time $\in (17.7; 131.0) \wedge$ standard bicarbonate $\in (23.8; 27.3) \wedge$ alfentanil administration $\in (0.0; 42.3)$
θ^F_{SDSD}	5	milrinone administration $\in (3.5; 635.4) \wedge$ cefazolin administration $\in (0.0; 92.5) \wedge$ prothrombin time $\in (14.8; 131.0)$
θ^F_{SDSD}	6	prothrombin time $\in (17.7; 131.0) \wedge$ base excess $\in (-0.7; 3.2) \wedge$ calcium ions $\in (1.1; 1.2)$
θ^F_{SDSD}	7	fibrinogen $\in (1.7; 7.7) \wedge$ standard bicarbonate $\in (23.8; 30.2) \wedge$ drains $\in (0.0; 480.0)$
θ^F_{SDSD}	8	prothrombin time $\in (17.7; 131.0) \wedge$ standard bicarbonate $\in (23.8; 27.3) \wedge$ calcium ions $\in (1.1; 1.2)$
θ^F_{SDSD}	9	prothrombin time $\in (17.7; 131.0) \wedge$ standard bicarbonate $\in (23.8; 27.3) \wedge$ sodium $\in (133.0; 141.0)$
θ^F_{SDSD}	10	prothrombin time $\in (17.7; 131.0) \wedge$ blood loss $\in (0.0; 958.8) \wedge$ age $\in (51; 74)$
θ^F_{SDSD}	11	prothrombin time $\in (17.7; 131.0) \wedge$ Ringer's lactate administration $\in (0.0; 4813.1) \wedge$ FIO2 $\in (0.2; 0.2)$
θ^F_{SDSD}	12	prothrombin time $\in (17.7; 131.0) \wedge$ blood loss $\in (0.0; 958.8) \wedge$ fluid loss $\in (0.0; 0.0)$
θ^F_{SDSD}	13	prothrombin time $\in (17.7; 131.0) \wedge$ blood loss $\in (0.0; 958.8) \wedge$ chloride $\in (105.0; 111.0)$
θ^F_{SDSD}	14	prothrombin time $\in (16.5; 131.0) \wedge$ standard bicarbonate $\in (23.7; 27.3) \wedge$ potassium $\in (4.0; 5.0)$
θ^F_{SDSD}	15	milrinone administration $\in (3.5; 635.4) \wedge$ cefazolin administration $\in (0.0; 92.5) \wedge$ FIO2 $\in (0.2; 0.5)$

Table 10: Top-15 subgroups discovered when seeking exceptional phenotype θ^F_{RMSSD}.

Phenotype	#	Description
θ^F_{RMSSD}	1	prothrombin time $\in (16.5; 131.0) \wedge$ standard bicarbonate $\in (23.7; 27.3) \wedge$ thrombin administration $\in (0.0; 0.0)$
θ^F_{RMSSD}	2	fibrinogen $\in (1.7; 7.7) \wedge$ base excess $\in (-0.6; 7.3) \wedge$ thrombin administration $\in (0.0; 0.0)$
θ^F_{RMSSD}	3	prothrombin time $\in (17.7; 131.0) \wedge$ standard bicarbonate $\in (23.8; 27.3) \wedge$ alfentanil administration $\in (0.0; 42.3)$
θ^F_{RMSSD}	4	prothrombin time $\in (17.7; 131.0) \wedge$ base excess $\in (-0.7; 3.2) \wedge$ calcium ions $\in (1.1; 1.2)$
θ^F_{RMSSD}	5	prothrombin time $\in (17.7; 131.0) \wedge$ cefazolin administration $\in (0.0; 90.3) \wedge$ FIO2 $\in (0.2; 0.2)$
θ^F_{RMSSD}	6	prothrombin time $\in (17.7; 131.0) \wedge$ standard bicarbonate $\in (23.8; 27.3) \wedge$ calcium ions $\in (1.1; 1.2)$
θ^F_{RMSSD}	7	prothrombin time $\in (16.5; 131.0) \wedge$ standard bicarbonate $\in (23.7; 27.3) \wedge$ chloride $\in (105.0; 110.0)$
θ^F_{RMSSD}	8	prothrombin time $\in (17.7; 131.0) \wedge$ standard bicarbonate $\in (23.8; 27.3) \wedge$ sodium $\in (133.0; 141.0)$
θ^F_{RMSSD}	9	prothrombin time $\in (17.7; 131.0) \wedge$ blood loss $\in (0.0; 958.8) \wedge$ age $\in (51; 74)$
θ^F_{RMSSD}	10	milrinone administration $\in (3.5; 635.4) \wedge$ cefazolin administration $\in (0.0; 92.5) \wedge$ prothrombin time $\in (14.8; 131.0)$
θ^F_{RMSSD}	11	prothrombin time $\in (17.7; 131.0) \wedge$ noradrenalin administration $\in (0.0; 24.6) \wedge$ fluid loss $\in (0.0; 0.0)$
θ^F_{RMSSD}	12	prothrombin time $\in (17.7; 131.0) \wedge$ blood loss $\in (0.0; 958.8) \wedge$ fluid loss $\in (0.0; 0.0)$
θ^F_{RMSSD}	13	prothrombin time $\in (17.7; 131.0) \wedge$ blood loss $\in (0.0; 958.8) \wedge$ chloride $\in (105.0; 111.0)$
θ^F_{RMSSD}	14	fibrinogen $\in (1.7; 7.7) \wedge$ standard bicarbonate $\in (23.8; 30.2) \wedge$ drains $\in (0.0; 480.0)$
θ^F_{RMSSD}	15	prothrombin time $\in (17.7; 131.0) \wedge$ base excess $\in (-0.7; 3.2) \wedge$ coronary artery bypass grafting = False

Table 11: Top-15 subgroups discovered when seeking exceptional phenotype θ^F_{SDSQ}.

Phenotype	#	Description
θ^F_{SDSQ}	1	prothrombin time $\in (17.7; 131.0) \wedge$ noradrenalin administration $\in (0.0; 24.6) \wedge$ fluid loss $\in (0.0; 0.0)$
θ^F_{SDSQ}	2	prothrombin time $\in (17.7; 131.0) \wedge$ standard bicarbonate $\in (23.8; 27.3) \wedge$ age $\in (51; 74)$
θ^F_{SDSQ}	3	prothrombin time $\in (17.7; 131.0) \wedge$ noradrenalin administration $\in (0.0; 24.6) \wedge$ taking pantoprazole = False
θ^F_{SDSQ}	4	prothrombin time $\in (16.5; 131.0) \wedge$ standard bicarbonate $\in (23.7; 27.3) \wedge$ thrombin administration $\in (0.0; 0.0)$
θ^F_{SDSQ}	5	prothrombin time $\in (17.7; 131.0) \wedge$ cefazolin administration $\in (0.0; 90.3) \wedge$ FIO2 $\in (0.2; 0.2)$
θ^F_{SDSQ}	6	prothrombin time $\in (17.7; 131.0) \wedge$ Ringer's lactate administration $\in (0.0; 4813.1) \wedge$ FIO2 $\in (0.2; 0.2)$
θ^F_{SDSQ}	7	prothrombin time $\in (17.7; 131.0) \wedge$ cell salvage administration $\in (0.0; 646.0) \wedge$ taking pantoprazole = False
θ^F_{SDSQ}	8	prothrombin time $\in (17.7; 131.0) \wedge$ blood loss $\in (0.0; 958.8) \wedge$ age $\in (51; 74)$
θ^F_{SDSQ}	9	prothrombin time $\in (17.7; 131.0) \wedge$ Ringer's lactate administration $\in (0.0; 4813.1) \wedge$ amiodarone administration $\in (0.0; 0.0)$
θ^F_{SDSQ}	10	prothrombin time $\in (17.7; 131.0) \wedge$ Ringer's lactate administration $\in (0.0; 4813.1) \wedge$ minimally invasive aortic valve replacement = False
θ^P_{SDSQ}	11	fibrinogen $\in (1.7; 7.7) \wedge$ age $\in (51; 65) \wedge$ base excess $\in (-1.1; 3.2)$
θ^F_{SDSQ}	12	prothrombin time $\in (17.7; 131.0) \wedge$ blood loss $\in (0.0; 958.8) \wedge$ fluid loss $\in (0.0; 0.0)$
θ^F_{SDSQ}	13	prothrombin time $\in (17.7; 131.0) \wedge$ Ringer's lactate administration $\in (0.0; 4813.1) \wedge$ heparin administration $\in (0.0; 11.0)$
θ^F_{SDSQ}	14	prothrombin time $\in (17.7; 131.0) \wedge$ cell salvage administration $\in (0.0; 646.0) \wedge$ taking atorvastatin = False
θ^F_{SDSQ}	15	prothrombin time $\in (17.7; 131.0) \wedge$ cell salvage administration $\in (0.0; 646.0) \wedge$ coronary artery bypass grafting = False

2 Experiment on Publicly Available Data

The experiments in [1, Section 6] are carried out on hospital data that is not available for public distribution. We repeat the experiments on a synthetic data set to allow reproduction. In the remainder of this section, we will refer to notations and terminology defined in [1, Section 3].

2.1 Synthetic Data

The data in this experiment combines synthetic and public data: synthetically generated Electronic Health Records (EHR), Electrocardiogram (ECG) signals from the CPSC 2021 database [4, 6], and corresponding Atrial Fibrillation (AF) indicators from the matching the signals from the database.

2.1.1 Electronic Health Records. The EHR characteristics function as the descriptors $(a_1, ..., a_M)$ of the Exceptional Model Mining (EMM) framework. We produce 57 realistic medical attributes, renamed to **a1, a2, ...** because of the sensitive nature of medical data. This process is performed at random while maintaining medical patterns. **(a1, ..., a4)** concern information, where the age is selected between 50 and 80 years, the sex is skewed towards men (this follows the trend in hospital data that more men get cardiac surgery [3]), the BMI is skewed towards healthy and slightly overweight patients, and the ASA score is randomly selected between 1 and 6, getting progressively worse. **(a5, ..., a7)** contains information about habits that randomly pick one of the options {None, Used, Uses}. **(a8, ..., a14)** has information on fluid loss during cardiac surgery, which has a 60% chance of being 0, and a 40% chance of being randomly sampled between 0 and 2500. Lastly, **(a15, ..., a57)** are medications that patients can take at home, the options are {Not taken, Taken} and they are sampled in an 80/20 ratio. Each combination of medical characteristics belongs to one patient.

2.1.2 Electrocardiogram Signals. The ECG signals form the basis for our targets $(\ell_1, ..., \ell_K)$. These targets are extracted from the ECG signals in the CPSC 2021 data set. It includes ECG signals from 105 patients with (49) and without (56) AF. The signals are stored in segments per patient and were originally divided into training sets I and II. However, we combined all of them into one data set. As the quantity of data segments was too large, it became infeasible to combine all segments to obtain an ECG per patient. Thus, we denoise and extract the features per segment, and concatenate the extracted features so that the final data set has one row per patient. One main difference between the private hospital data and this public data set is the sampling frequency, which is 500 Hz and 200 Hz respectively. The effect on denoising is negligible, but the R-peak correction is altered to fit the frequency. The extra denoising on the SQ-interval that originally contained two weighted average filters now only includes the first one ($\sigma = 20, M = 10$) to prevent oversmoothing (cf. [1, Section 5]). Ultimately, we successfully extract the ECG features of 97 of the 105 patients.

2.1.3 Atrial Fibrillation Indicators. A binary indicator (b_1) functions as the *evaluator*. These labels in the ECG data set indicate whether the patient does or does not experience AF in the recording.

Table 12: Top-1 subgroups discovered when seeking all eleven exceptional phenotypes on the publicly available data.

Phenotype	Description	φ	% AF
θ_{SDSD}	Using **a6** $\wedge$ **a13** $\in [1172; 1816]$	1.94	85.7
θ_{RMSSD}	Using **a6** $\wedge$ **a13** $\in [1172; 1816]$	3.02	85.7
θ_{SDRR}	**a14** $\in [881; 2332] \wedge$ **a54** is not taken $\wedge$ **a38** is not taken	1.72	83.3
θ_P	**a46** is taken $\wedge$ **a10** $\in [0; 0] \wedge$ **a11** $\in [0; 273]$	1.73	100.0
θ_F	**a43** is taken $\wedge$ **a3** is underweight	0.14	80.0
θ^P_{SDSD}	**a48** is taken $\wedge$ **a2** is male $\wedge$ **a9** $\in [221; 2256]$	2.58	87.5
θ^P_{RMSSD}	**a43** taken $\wedge$ **a12** $\in [84; 2463] \wedge$ **a34** is not taken	12.26	80.0
θ^P_{SDSQ}	**a43** taken $\wedge$ **a13** $\in [0; 0] \wedge$ **a47** is not taken	8.88	80.0
θ^F_{SDSD}	**a48** taken $\wedge$ **a2** is male $\wedge$ **a9** $\in [221; 2256]$	2.82	87.5
θ^F_{RMSSD}	**a43** taken $\wedge$ **a12** $\in [84; 2463] \wedge$ **a34** is not taken	12.26	80.0
θ^F_{SDSQ}	**a43** is taken $\wedge$ **a38** is not taken $\wedge$ **a47** is not taken	3.08	83.3

2.2 Results

For each of the eleven experiments, we deploy Exceptional Model Mining (EMM) on the descriptors and targets. One description is generated per experiment and described in Table 12. These descriptions do not contain sensitive information, as their sole purpose is reproducibility. The processed data is publicly available on GitHub[1]; it can also be processed anew. However, keep in mind that this takes over two hours. When deployed, a new ECG instance can be easily added to the processed list of quality measures and AF complications. Adding this takes about five minutes maximum. The beam search on these processed lists takes only 15 seconds.

References

[1] L. van den Biggelaar, R.M. Schouten, A. de Bie, R.A. Bouwman, and W. Duivesteijn. 2025. Characterizing the Risk of Atrial Fibrillation in Cardiac Patients with Exceptional Electrocardiogram Phenotypes. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2025).* https://doi.org/10.1145/3711896.3737200 , to appear.

[2] W. Duivesteijn, A. Feelders, and A. Knobbe. 2016. Exceptional model mining: Supervised descriptive local pattern mining with complex target concepts. *Data Mining and Knowledge Discovery* 30 (2016), 47–98. https://doi.org/10.1007/s10618-015-0403-4

[3] A. Ehdaie, E. Cingolani, M. Shehata, X. Wang, A.B. Curtis, and S.S. Chugh. 2018. Sex differences in cardiac arrhythmias: Clinical and research implications. *Circulation: Arrhythmia and Electrophysiology* 11, 3 (2018), e005680. https://doi.org/10.1161/CIRCEP.117.005680

[1] https://github.com/liekevandenbiggelaar/EAFM; our implementation builds upon the pseudocode from [2, Algorithm 1] and the implementation from [5].

[4] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.-K. Peng, and H.E. Stanley. 2000. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* 101, 23 (2000), e215–e220. https://doi.org/10.1161/01.CIR.101.23.e215

[5] R.M. Schouten, W. Duivesteijn, and M. Pechenizkiy. 2022. Exceptional model mining for repeated cross-sectional data (EMM-RCS). In *Proceedings of the 2022 SIAM International Conference on Data Mining (SDM)*. SIAM, 585–593. https://doi.org/10.1137/1.9781611977172.66

[6] X. Wang, C. Ma, Z. Zhang, H. Gao, G. Clifford, and C. Liu. 2021. Paroxysmal atrial fibrillation events detection from dynamic ECG recordings: The 4th China Physiological Signal Challenge 2021. data retrieved from Physionet, https://physionet.org/content/cpsc2021/1.0.0/.